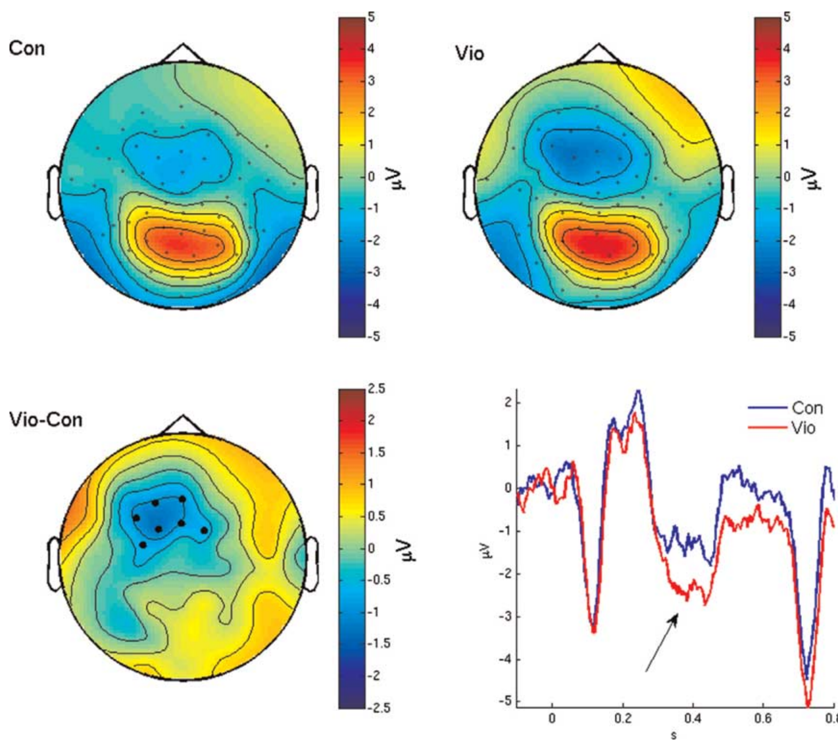
SUPPLEMENTARY SECTION 1

It is often the case that repeated-measures analysis of variance (ANOVA) is used to provide statistical support for inferences made about ERP patterns. This section describes an ANOVA analysis of the violation-control contrast reported in the main text of the paper, for purposes of comparison to the reported clustering and randomization approach (Maris & Oostenveld, 2007).

For this analysis, the 61 channel array was divided into quadrants by taking all electrodes (excluding the midline) in each of the left frontal (7, 17:19, 32:34, 48:50, 60:61), right frontal (3, 9:11, 21:23, 36:38, 51:52), right posterior (4, 12:13, 24:27, 39:42, 53:55), and left posterior (6, 15:16, 28:31, 44:47, 57:59) electrode groups indicated in Supplementary Figure 3. Two average responses were evaluated, corresponding to the average voltage in the time windows of 0.3 to 0.5 *s* (LAN time window) and 0.5 to 0.9 *s* (P600 time window).

The ANOVA for both time windows included the repeated measures factors Hemisphere (left, right), Direction (anterior, posterior), Condition (violation, control), and Participant (subjects 1 to 20). Neither ANOVA indicated a significant main effect of Condition, or interactions of Condition with Hemisphere or Direction. For completeness, the three-way interaction of Condition*Direction*Hemisphere for the LAN effect was $F(1, 19) = 1.2545$, $p=.28$, $MSE=0.0598$, and the two-way Condition*Hemisphere interaction test was $F(1, 19) = 2.325$, $p=.14$, $MSE=0.5504$. The main effect of Condition was $F(1, 19)$ was less than 1. For the P600 contrast, the two-way Condition*Direction interaction was $F(1, 19) = 3.356$, $p=.08$, $MSE=1.2876$, and the main effect of Condition also had an F-ratio less than 1.

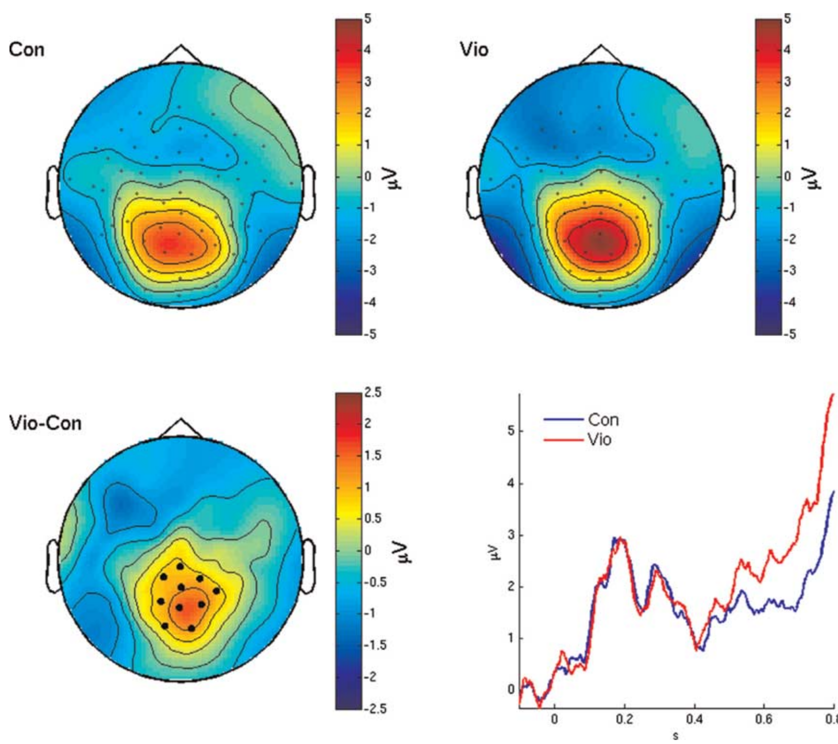
The results of the clustering and randomization tests reported in the main text indicated statistical support for a difference between violation and control conditions for each of these electrode groups. The difference between the ANOVA and the randomization tests is most likely due to the averaging over electrodes in the ANOVA analysis. Without an *a priori* selection of electrodes for the analysis, the ANOVA analysis averages across those channels that have a larger difference and those that do not. The purpose of the randomization tests was to find the channels that were sensitive to these contrasts.



Supplementary Figure 1. Average evoked potential for the control condition (Con), declension violation (Vio), difference topography (Vio-Con), and the time course of the average ERPs for the LAN effect. Black dots plotted on the difference topography indicate the electrodes showing a statistically significant contrast between violation and control. All topography plots show the average response (or response difference) in the time window 0.3 to 0.5 *s*. The ERP traces show the average response at these significant electrodes.

Finally, Supplementary Figure 4 shows a overlay of the average left frontal and posterior electrode groups for the LAN/P600 effects, the Declension*Preposition interaction effect, and the Declension*Preposition*$I_{x|w}$ interaction effect. Recall that the posterior electrode group showed a significant Declension*Preposition interaction effect (see Figure 3q), while both the frontal and posterior groups showed significant Declension*Preposition*$I_{x|w}$ interactions (see Figure 3i,s). Supplementary Figure 4 indicates for the frontal response that while there is a general trend for the interaction effect to follow a similar time course as the LAN effect, the earliest phase of the LAN effect (at approximately 0.3 s) is not present in the interaction. Note that statistical support for the frontal three-way interaction was only found late within the response interval. For the posterior response, the three-way interaction appears to occur approximately 0.1 s later than the P600, while the two-way Declension*Preposition interaction occurs considerably earlier.

Although the late three-way interaction observed here is presumably not the reflection of a reanalysis process (because no error was present), perhaps it might be considered a learning process analogous to the proposal by Barber and Carreiras (2005) described in the introduction, in which stimuli with high information value are stored in memory. However, it should also be pointed out that Barber and Carreiras (2005) attributed the late-phase P600 they observed for Spanish noun gender violations to a lexical reanalysis process, and this is unlike the stimulus presented here, which is a grammatically correct inflected adjective.



Supplementary Figure 2. Average evoked potential for the control condition (Con), declension violation (Vio), difference topography (Vio-Con), and the time course of the average ERPs for the P600 effect. Black dots plotted on the difference topography indicate the electrodes showing a statistically significant contrast between violation and control. All topography plots show the average response (or response difference) in the time window 0.5 to 0.9 *s*. The ERP traces show the average response at these significant electrodes.

SUPPLEMENTARY SECTION 2

The experimental design reported in the main text included factors for declension (strong, weak), preposition (dative, accusative), as well as gender (masculine, neuter). This section describes wavelet-regression results for the factor gender. As Table 3 indicates, the design does not include a full crossing of these three factors. We therefore only included in the analysis the two-way interactions between declension and gender, on the one hand, and preposition and gender, on the other. For the left frontal response, the wavelet regression approach indicated relatively late main effects of declension (Supplementary Figure 5b) and preposition (Supplementary Figure 5c), but no main effect of gender, nor any interactions of gender with declension or preposition (Supplementary Figure 5d–f). For the posterior electrode group, there was a similar pattern (e.g., Supplementary Figure 5h,i for the main effects, and Supplementary Figure 5j–l for the lack of interactions with gender).

SUPPLEMENTARY SECTION 3

In the analysis of the adjectival response, the baseline interval was the interval before the onset of the adjective. As pointed out by an anonymous reviewer, there is a potential confound for the analysis because the phrases are not the same before the adjective. In the strong condition, the interval of time before the adjective onset includes the response to the previous preposition. In the weak condition, this interval includes the response to the previous determiner, which itself follows a preposition. The response to the weak condition might therefore already differ from the strong condition, even before the onset of the adjective. If this difference occurs in the baseline interval, then the pre-existing difference should appear as a constant effect in the adjectival response.

Besides this word category difference (preposition versus determiner), there is also another morpho-syntactic difference. In the weak condition, the determiner provides morphosyntactic information that is related to the form of the upcoming adjective (see Introduction in the main text). Participants might, in principle, predict the suffix form of the correctly inflected phrases prior to the onset of the adjective.

This supplementary section provides an analysis of the ERPs to a longer time series including both the word preceding the adjective, as well as the adjective, to address this property of the phrases.

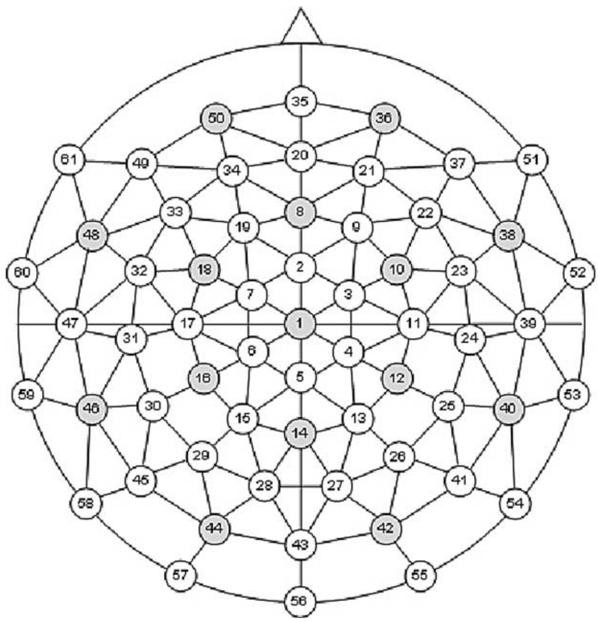
The time interval from −0.7 before and 0.9 s after the adjective was analysed using the same factors as the adjective-only analysis reported in the main text. The ERPs for the strong and weak conditions were baselined with respect to the interval from −0.1 s to 0.0 s before the onset of the preposition or determiner, respectively. Supplementary Figure 6 shows the average response and the effects of the main experimental factors on this time series for the left frontal (Supplementary Figure 6a–j) and posterior (Supplementary Figure 6k–t) electrode groups. Note that the response in the strong condition is the response to the preposition and adjective, while the response in the weak condition is the response to the determiner and the adjective. In the plotted time scale, the onset of the preposition/determiner occurs at time −0.6 s and the onset of the adjective occurs at time 0 s.

The posterior response showed statistical effects that were similar to the original analysis except for the main effect of Declension, and there were no substantial effects in the time interval preceding the adjective. The main effect of Declension was a sustained response beginning at approximately 0.1 s after the adjective onset, lasting until approximately 0.7 s. The interaction between Declension and Preposition was less robust than in the main analysis, as was the higher order interaction between Declension, Preposition, and $I_{x|w}$.

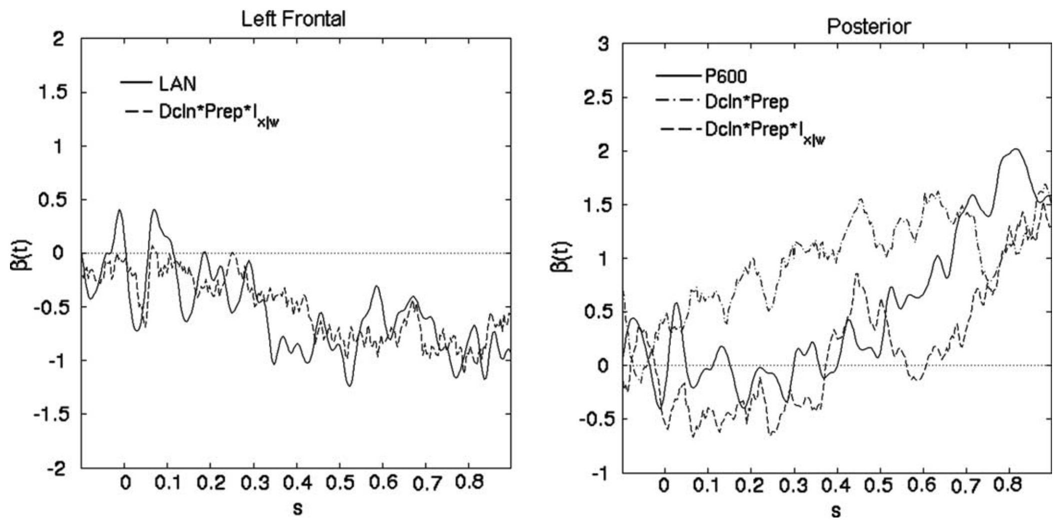
For the left frontal response, there was a sustained deviation from the average for the factor Declension beginning at approximately 0.2 s until the end of the response interval (Supplementary Figure 6f). For the interaction between Declension and Preposition, there were several brief modulations of this main effect in the adjective-preceding time window. There was little evidence for the three-way interaction between Declension, Preposition, and $I_{x|w}$ that was present in the analysis presented in the main text, although there was a very brief modulation late in the response window, consistent with that analysis.

The critical question is whether there is a pre-existing difference in the pre-adjectival response in the baseline interval for the adjectival response. The ISI between word onsets in this experiment was 0.6 s, which means that the baseline interval for the adjectival response would be represented in the 0.5 to 0.6 s time interval in the current analysis of the preposition/determiner response. For this interval in the left frontal response, the main effect for the contrast between strong and weak was clearly present (Supplementary Figure 6f). It was not present for the posterior response (Supplementary Figure 6p). The interaction between Declension and Preposition was not apparent in this interval, for either the left frontal (Supplementary Figure 6g) or posterior response (Supplementary Figure 6q). For none of the interactions is there an effect apparent in this interval.

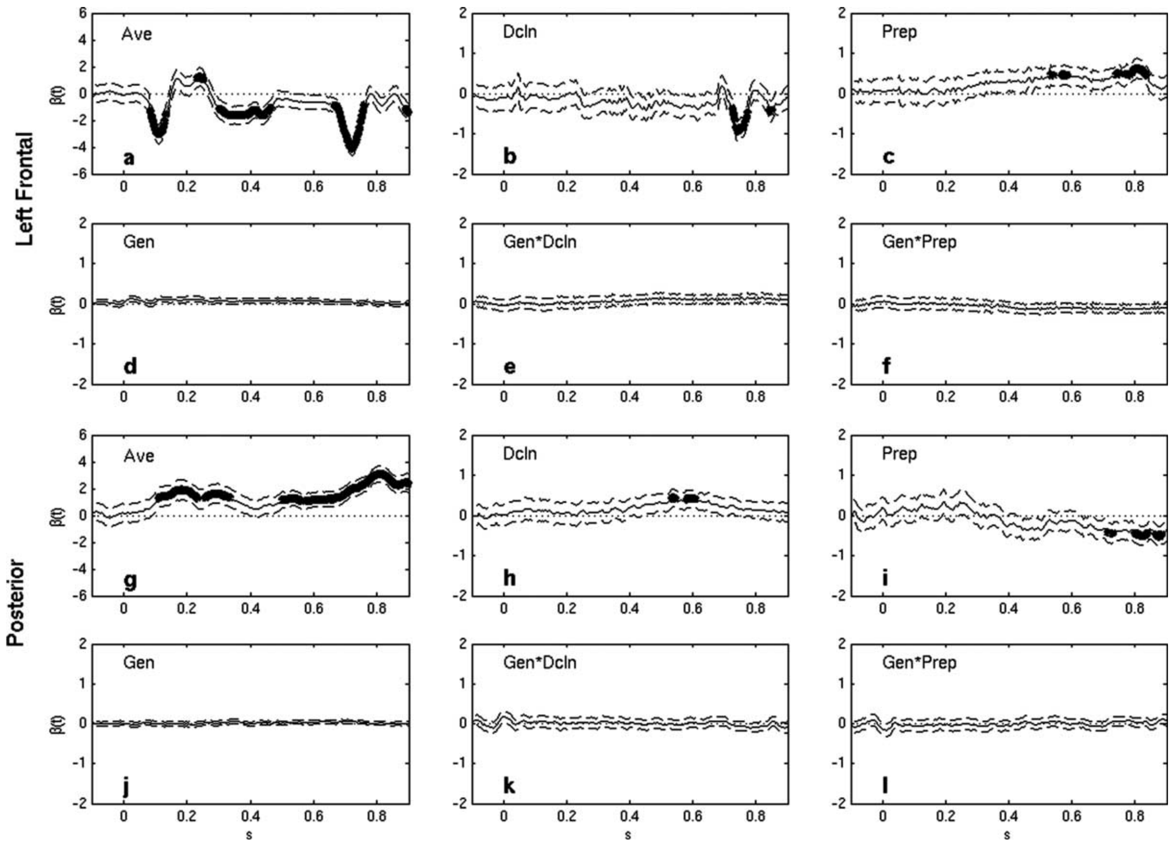
These results suggest that in the left frontal response, the main effect of the Declension contrast at the pre-adjectival position might have spilled over to the adjectival response. There is no indication of an overlapping component for the posterior response. As expected, there was no relationship between the information quantity regressors and the pre-adjectival response. The results show clear evidence of activity prior to the onset of the adjective, and that for the frontal response, this activity was different for the strong and weak conditions. However, because none of the higher order interactions indicated activity in the baselining interval, we would suggest that the results reported in the main text are not due (solely) to activity in the baseline period.



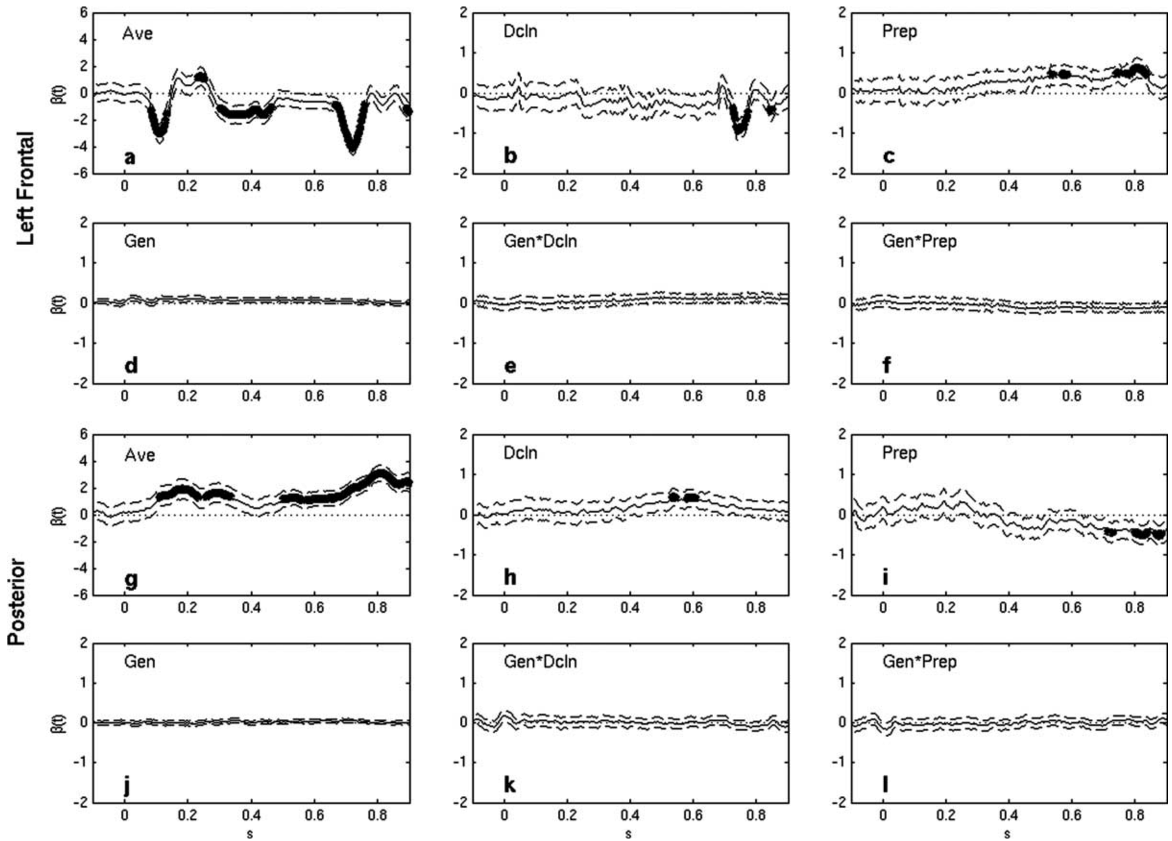
Supplementary Figure 3. Electrode arrangement for the EEG recordings. Approximate 10-20 locations are shown in grey.



Supplementary Figure 4. Coefficients (functional beta weights) for the left frontal and posterior Declension*Preposition and Declension*Preposition*$I_{x|w}$ interaction effects superimposed on the LAN (left frontal) and P600 (posterior) effects.



Supplementary Figure 5. Coefficients (functional beta weights) for the left frontal (a–f) and posterior (g–l) electrode groups for the effects of the experimental factors declension (*Dcln*), preposition (*Prep*), and gender (*Gen*) on the response to the adjective (see supplementary text for a description).



Supplementary Figure 6. Coefficients (functional beta weights) for the left frontal (a–j) and posterior (k–t) electrode groups for the effects of the experimental factors declension (*Dcln*), preposition (*Prep*), and their interactions with various information quantity regressors on the response to the word before the adjective as well as the adjective (see main text for a description).